



# Software Technical Committee Building Software for Software Development Analytics

January 2018

Jesus M. Gonzalez-Barahona, Bitergia / URJC  
Harish Pillay, Red Hat  
Sean Goggins, University of Missouri



# Why?

We need to understand how software is being developed:

- Open Development Analytics
- Project Health

Activity, community, processes, signals, values, and goals.

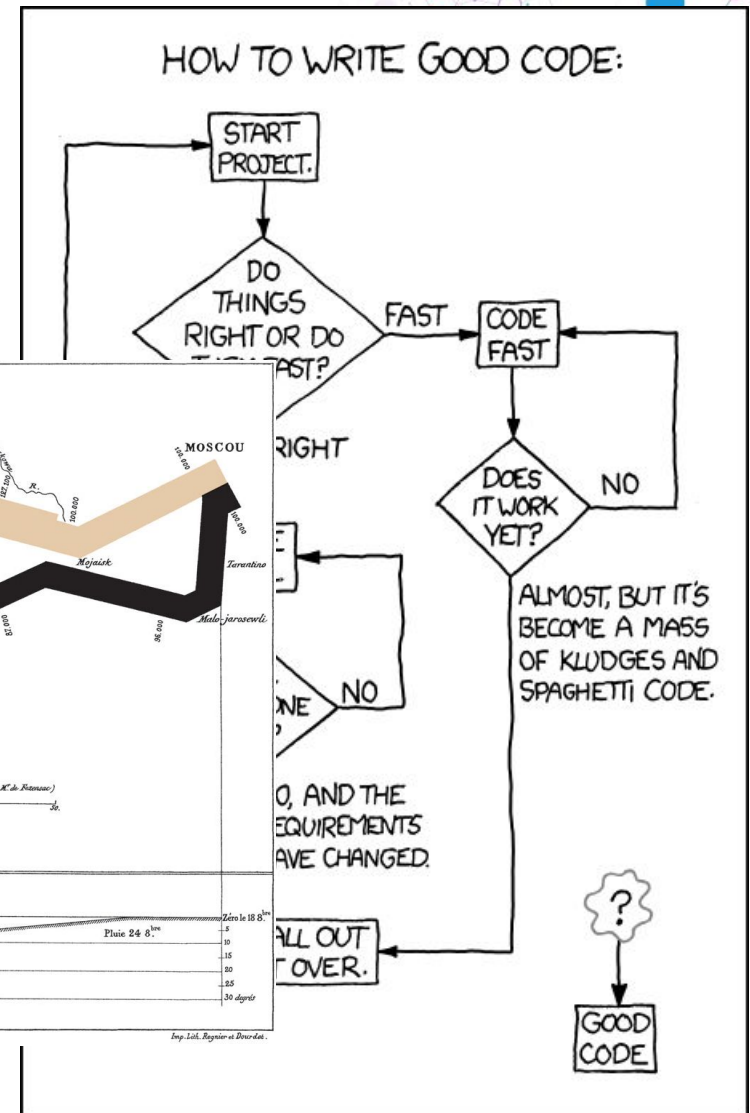
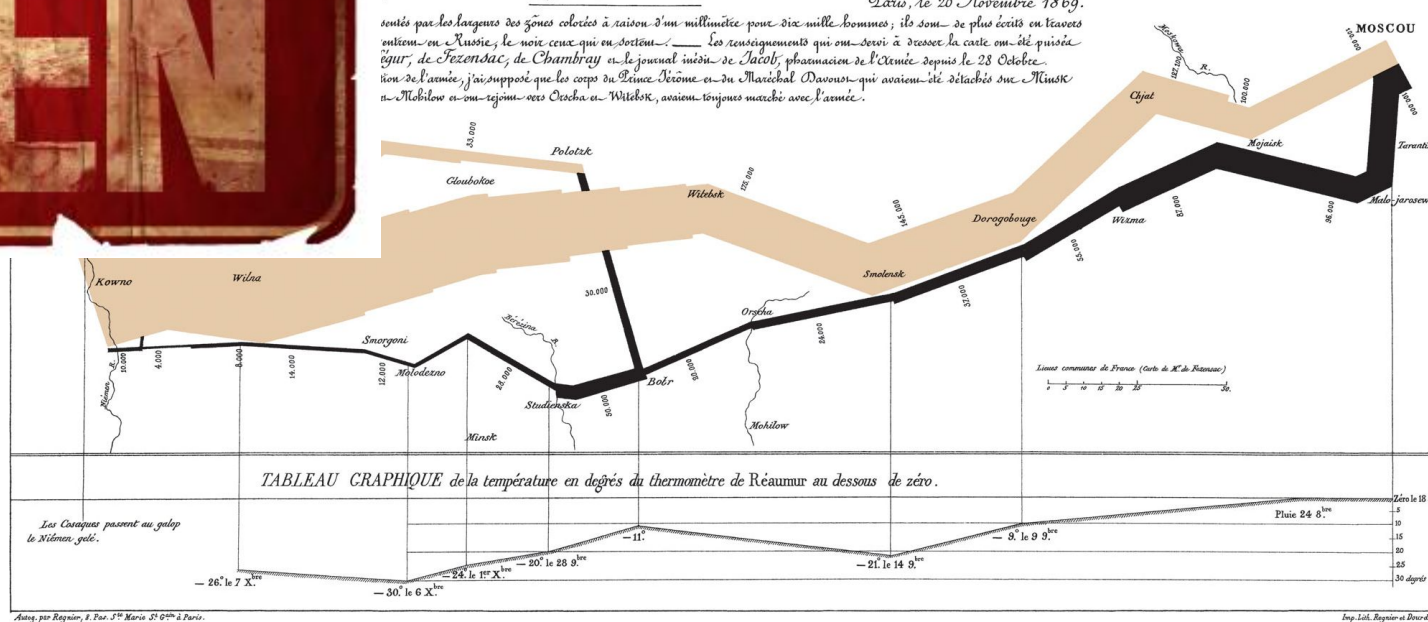


# Open Development Analytics



études successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.  
 par M. Marné, Inspecteur Général des Ponts et Chaussées en retraite. Paris, le 20 Novembre 1869.

ontés par les larges zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en lettres entières en Russie, le noir ceux qui ont servi à dresser la carte ont été puisés à la source, de Tcherniac, de Chambray et le journal intime de Jacob, pharmacien de l'Armée depuis le 28 Octobre. Pour de l'année, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Mabilow et ont rejoint vers Orscha et Witebsk, avaient toujours marché avec l'armée.





# Open Development Analytics

A step beyond in  
project  
transparency



*From “show me the code”  
to “show me the numbers”*



CHA<sup>O</sup>SS

[chaoss.community](http://chaoss.community)

CHA<sup>O</sup>SS

# Mission

*Produce integrated, open source software for analyzing software development, and definition of standards and models used in that software in specific use cases*

*Establish implementation-agnostic metrics for measuring community activity, contributions, and health*

*Optionally produce standardized metric exchange formats, detailed use cases, models, or recommendations to analyze specific issues in the industry/OSS world*



# How?

*“Meeting point for people defining & using software metrics  
and  
people implementing analytics for software development”*



# Founding Members





# How? Structure



## Two Committees:

### Metrics

Implementation-agnostic  
community development  
metrics

### Software

Integrated FOSS tools for  
software development  
analytics

# CHA<sup>O</sup>SS



## Software TC

[chaoss.community](https://chaoss.community)

CHA<sup>O</sup>SS

# Prospector

credit

GHDData

CHA<sup>OS</sup>SS



# Prospector



- Open source projects vary greatly in strength, significance, vibrancy and influence
- No simple way to evaluate or compare projects objectively, other than through individual experts
- Risk of committing to declining projects or missing out on thriving ones
- Open Source projects are not always openly trackable

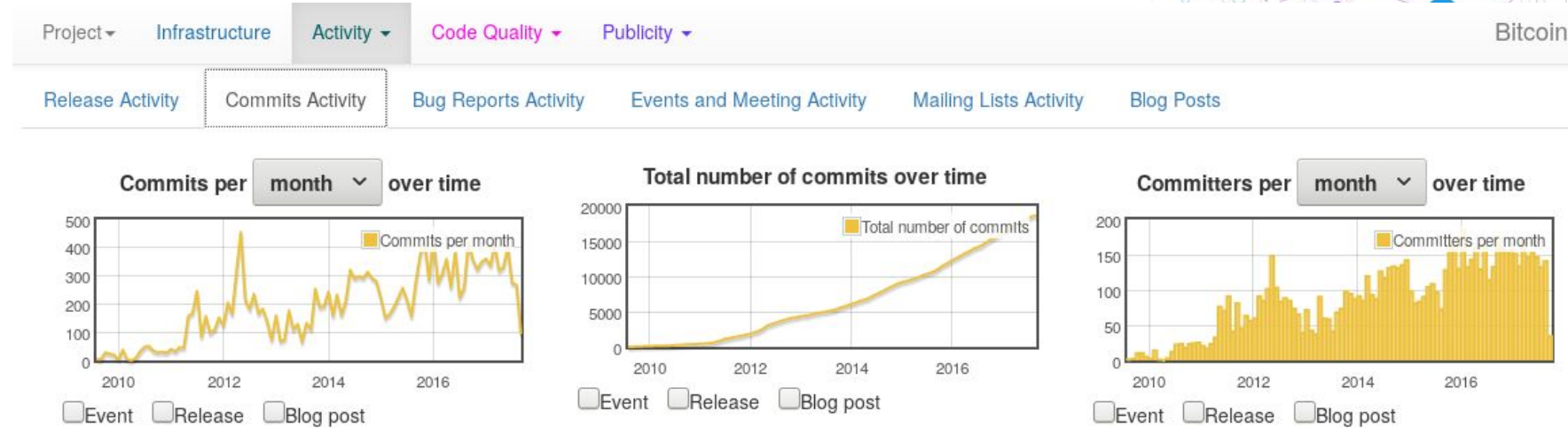
# Prospector - How?



- Providing an objective, consistent and repeatable set of metrics of projects for success, sustainability and vibrancy.
- These can then coherently help assess and track continuously open source projects, which in turn would help drive the evolution of projects



# Prospector



Metric datapoints:

- **Percentage of committers by dominant domain name: 37.01%**  
Rationale: If more than 50% of COMMITTERS are from one domain (via email ID) it is dominated by one set of people. Suggested target is to have it less than 35%.
- **Percentage of commits by dominant domain name: 57.24%**  
Rule: If more than 50% of CODE COMMITS are from one domain (via email ID) it is dominated by one set of people. Suggested target is to have it less than 35%.
- **Unique email address domains: 345 domains**  
How many unique domains are represented?
- **Unique committers: 846 committers**  
How many UNIQUE COMMITTERS are represented?

Number of commits over last:

- **3 months:** 850 commits from 81 contributors
- **6 months:** 1,985 commits from 131 contributors
- **9 months:** 2,984 commits from 164 contributors
- **12 months:** 4,176 commits from 233 contributors
- **24 months:** 8,028 commits from 412 contributors

# cregit



Framework to create evolutionary views of source code stored in a git repository,

Allow the summarizing of contributions at token, function, or file level.

Current support for C, C++, Java, go, and python

[github.com/cregit](https://github.com/cregit)



# Credit: improving accuracy of git-blame



- **Git-blame tracks changed lines, not tokens**
  - Last person who modified part of a line, becomes “contributor” of the entire line
  - Credit is capable of tracking the contributor of each token in a line
- **In Linux:**
  - blame per line is accurate in 75%
  - blame per token (using credit) is accurate 95%
  - Results based on statistical sampling and manual analysis, with 95% reliability with +/-5% of error
- **Currently in use by the Linux Kernel**

[credit.linuxsources.org](https://credit.linuxsources.org)

# cregit



```
cpu_notifier_register_begin();

if (create_hash_tables()) {
    err = -ENOMEM;
    goto out;
}

entry = proc_create("profile", S_IWUSR | S_IRUGO,
                    NULL, &proc_profile_operations);
if (!entry)
    goto out;
proc_set_size(entry, (1 + prof_len) * sizeof(atomic_t));
__hotcpu_notifier(profile_cpu_callback, 0);

out:
cpu_notifier_register_done();
return err;
}
```

## Contributors

<code>william lee irwin iii</code>	william lee irwin iii	58	58.00%
<code>srivatsa s. bhat</code>	srivatsa s. bhat	26	26.00%
<code>paolo ciarrocchi</code>	paolo ciarrocchi	5	5.00%
<code>david howells</code>	david howells	4	4.00%
<code>denis v. lunev</code>	denis v. lunev	4	4.00%
<code>al viro</code>	al viro	2	2.00%
<code>dave hansen</code>	dave hansen	1	1.00%

# GrimoireLab



## Software Development Analytics Toolset

- Retrieval from +30 data sources
  - Storage of all metadata (ElasticSearch)
  - Computing of interesting metrics
  - Visualization
  - Reports
- CHA<sup>SS</sup>



[grimoirelab.github.io](https://grimoirelab.github.io)



# How? Software Committee GrimoireLab



## PERCEVAL

*Sir Perceval* goes on the quest to retrieve and gather data from git, GitHub, Bugzilla, JIRA, Gerrit, mbox, pipermail, StackExchange, Discourse, etc.



## SORTING HAT

Tool to manage people identities information, to merge multiple person identities across different data sources, enrich profile information with affiliation information, etc.



## GRIMOIREELK

*Playground* for testing the whole set of tools as a platform for software development analytics. It's a prototype of the *Grimoire Open Development Analytics* platform.



## ARTHUR

*King Arthur* commands his loyal knight *Sir Perceval* managing the tasks to retrieve data for analysis. It manages data incremental update, parallel downloading, etc.



## KIBITER

Custom fork of Kibana to work on new ideas for metrics & data visualization to be used by GrimoireLab Panels.



## PANELS

Set of pre-defined widgets and dashboard templates to visualize Elasticsearch indexes generated by GrimoireELK with KiBiter.



## MORDRED




Docker container to help you deploying the Grimoire Lab using a set of configuration files.

Example of a dashboard:

[opnfv.biterg.io](http://opnfv.biterg.io)

# How? Software Committee GrimoireLab





**Bitergia**  
analytics

Create your own dashboard at [Cauldron.io](#) [Share](#) [Options](#) [Last 2 years](#)

Overview

Welcome!

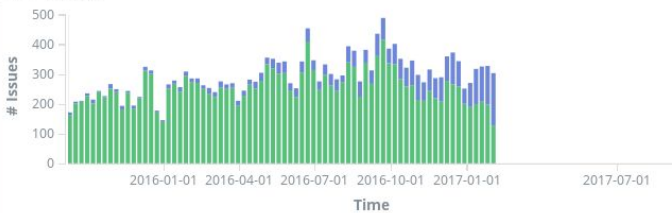
Click on cells, bars, slices, drag & drop on widgets or set a time range on the top right corner to filter data. Filters appear above as small stickers. [More info](#)

Top Submitters (Issues & PRs)

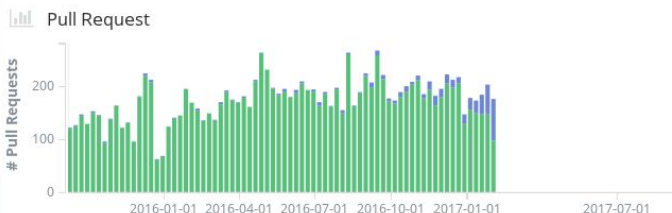
Avatar	Name	Profile	Issues	PRs
	Jeremy Elbourn	<a href="#">jelbourn</a>	186	285
	Jeff Cross	<a href="#">jeffcross</a>	155	116
	PatrickJS	<a href="#">gdi2290</a>	153	104
	Ward Bell	<a href="#">wardbell</a>	107	433
	tigbro	<a href="#">tbosch</a>	103	232

Export: [Raw](#) [Formatted](#)


Issues



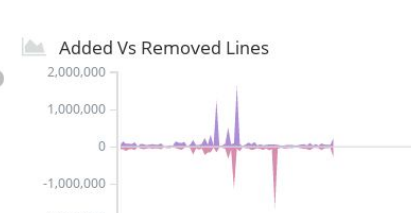
Pull Request



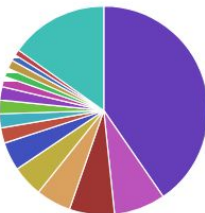
Commits



Added Vs Removed Lines



Domains



Top Repositories

Repositories	Commits	Issues	Pull Request
<a href="#">angular/angular</a>	4,334	10,155	3388
<a href="#">angular/angular.io</a>	2,234	2,985	2031
<a href="#">angular/material</a>	2,046	5,868	1630
<a href="#">angular/angular.js</a>	2,286	2,866	1304
<a href="#">angular/material2</a>	1,059	2,920	1293
<a href="#">angular/angular-cli</a>	1,036	4,361	1285

Overview

Git

Issues

Issues Timing

Pull Requests

PRs Timing

Backlog

About

Powered by [Grimoire Lab](#)

Provided by [Bitergia](#)

[Send us feedback](#)

[Collapse](#)

# GHData

Python library and web service for GitHub  
Health and Sustainability metrics

Uses data from the GHTorrent database and  
other sources

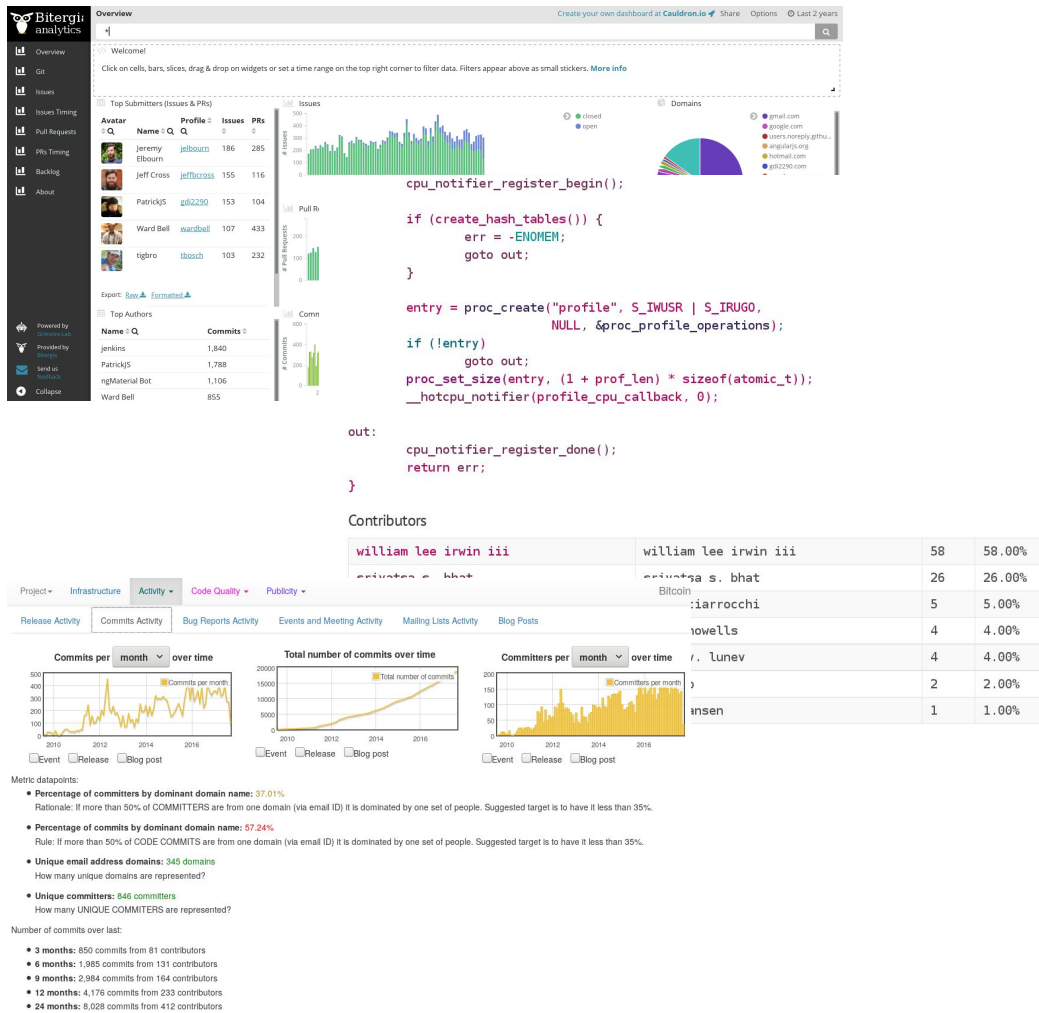
[osshealth.github.io/ghdata](https://osshealth.github.io/ghdata)

[github.com/OSSHealth/ghdata](https://github.com/OSSHealth/ghdata)



# How? Software Committee

## Live demos, proofs of concept



GrimoireLab:

[opnfv.biterg.io](http://opnfv.biterg.io)

[cauldron.io](http://cauldron.io)

Prospector:

[prospector.bitergia.net](http://prospector.bitergia.net)

Credit:

[credit.linuxsources.org/](http://credit.linuxsources.org/)

GHData:

[github.com/OSSHealth/ghdata](https://github.com/OSSHealth/ghdata)

# Building a Community



You're welcome to join!!



# Building a Community

## How to Join

Mailing Lists

Periodic Meetings

Weekly Hangouts

IRC Channels

Issues

Pull requests



[cnaoss.community](https://cnaoss.community)

Let's work together  
to extract knowledge  
from CHAOSS!!!!



[chaoss.community](https://chaoss.community)

CHAOSS

# License & Credits

(c) 2017-18 CHAOSS. Some rights reserved. This presentation is distributed under the “Attribution-ShareAlike 4.0” license, by Creative Commons, available at [creativecommons.org/licenses/by-sa/4.0/](https://creativecommons.org/licenses/by-sa/4.0/)



- “Napoleon’s Russian campaign of 1812”, Original by Charles Minard  
License: Public domain  
[en.wikipedia.org/wiki/Charles\\_Joseph\\_Minard#/media/File:Minard.png](https://en.wikipedia.org/wiki/Charles_Joseph_Minard#/media/File:Minard.png)
- “Aged Come In We’re Open”, Picture by Czarina Alegre in Flickr  
License: Creative Commons Attribution 2.0  
[flic.kr/p/fjGamh](https://flic.kr/p/fjGamh)
- “Good code”, Comic by Randall Munroe, XKCD 844  
License: Creative Commons Attribution-NonCommercial 2.5  
[xkcd.com/844/](https://xkcd.com/844/)